DOCUMENT RESUME

ED 137 186 SO 009 903

AUTHOR Napier, John D.
TITLE Studies of the Ability of Preservice Social Studies Teachers to Stage Score Moral Thought Statements.
PUB DATE Apr 77
NOTE 18p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, New York, April 3-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Developmental Psychology; Educational Experiments; *Educational Research; Evaluation; Higher Education; Measurement Instruments; *Moral Development; *Preservice Education; Scoring; *Social Studies; Tables (Data); Teacher Participation
IDENTIFIERS *Kohlberg (Lawrence); *Kohlberg Moral Development Test

ABSTRACT

The report describes two experiments involving the ability of preservice social studies teachers to stage score moral thought statements. Stage scoring is defined as keeping a record of statements in accordance with the stages of moral development originated by psychologist Lawrence Kohlberg. The two experiments involved the use of three stage scoring rater guides. They were designed to help teachers overcome content influence in order to stage score correctly on the basis of the structure of moral thought. The procedure for the first experiment was to randomly assign 32 preservice social studies teachers enrolled in a required methods course to two treatment groups. One group was given Kohlberg's sentence rater guide and the other was given his global rater guide. In the second experiment, 40 preservice teachers in the same course the following semester were assigned to similar treatment groups, one of which used an additional, updated, global rater manual. For both experiments, the preservice teachers were given information on Kohlberg's moral education program and instruction on how to use the rating guides. Findings indicate that none of the stage scoring rater guides aided teachers to overcome content influence and that, therefore, teachers should refrain from stage scoring until further research indicates which factors cause successful use of the stage scoring system. References related to teacher training, moral development, and assessing the moral reasoning of students are included. (Author/DB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED137180

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

STUDIES OF THE ABILITY OF PRESERVICE SOCIAL STUDIES TEACHERS TO STAGE SCORE MORAL THOUGHT STATEMENTS

John D. Napier
University of Georgia

ABSTRACT

The question of which Aspect stage scoring rater guide aided teachers to stage score correctly and overcome content influence while stage scoring was studied in two experiments. For both experiments, perservice teachers were given information on Kohlberg's theory of and education program for moral development prior to being randomly assigned to two treatments. Results from both experiments indicated that none of the treatments aided teachers to overcome content influence while stage scoring. Therefore, until research indicates what factors cause successful use of Kohlberg's stage scoring system, teachers should refrain from stage scoring.

In recent years there has been much discussion about using the cognitive-developmental approach to moral education in social studies education curricula. One of the teacher activities implied in most of the programs based on the theory of moral development originally researched by Kohlberg[1] was stage scoring moral thought statements. The review of "Kohlbergian" programs by Rest (1974) indicated that there were three purposes for stage scoring moral thoughts of students. First, the before and after instruction stage scoring was used to evaluate changes in the moral thoughts of students. Second, the before instruction stage scoring was used to arrange students into discussion groups containing different stage types. Third, the before and during instruction stage scoring was made so teachers knew the stages of students' moral thoughts in order to supply "proper" retorts during class discussions.

SO 009 903

A paper presented to the Special Interest Group/Research in Social Studies Education at the annual meeting of the American Educational Research Association, New York, 1977.

The use of stage scoring for evaluation purposes is apparent. However, the use of stage scoring to form discussion groups and supply "proper" retorts needs further elaboration. The reason given for forming discussion groups of different stage type students and having a teacher supply "proper" retorts rested on interpretations of the research of moral development and moral education. Interpretations by Kohlberg and his associates of studies conducted by people like Blatt, Rest, and Turiel (Blatt & Kohlberg, 1975; Rest, 1973; Rest, Kohlberg & Turiel, 1969; Turiel, 1969) concluded that individuals needed to be exposed to moral thought one stage above their own stage, termed +1 modeling, before moving to the next higher stage. As a result of these interpretations, teachers were to arrange students into small groups where +1 models were present for most members. They also were to stage score instantaneously moral thought statements made by students during discussions in order to supply "proper" retorts (i.e., +1 models). "Proper" retorts to a student during discussion could be handled by a teacher selecting a second student to respond at a +1 stage of moral thought, or by the teacher responding at a +1 stage of moral thought.

Recent literature has differed regarding whether teachers needed to and could stage score moral thought statements for all three of the previously mentioned purposes. Galbraith and Jones (1976) stated that teachers did not have to stage score moral thought statements during class discussions because the natural mixture of different moral stage types in a given class would automatically expose students to +1 models. However, they stated that teachers could stage score moral thought statements before and after instruction with the aid of a rater guide. On the other hand, Fenton and Kohlberg (1976a, 1976b) stated that with practice teachers could stage score moral thought statements before and during class discussions in order to supply +1 models. But, Fenton and Kohlberg stated that teachers cannot stage score before and after instruction for the purpose of evaluating changes in students' moral stages because Kohlberg's measurement system was too complicated for teachers to

use.[2] Despite the confusion on if and when teachers can and should stage score moral thought statements, stage scoring was a teacher activity for all four authors sometime during the implementation of the cognitive-developmental approach to moral education.

Unfortunately, a previous study (Napier, 1976) indicated that elementary school teachers were unable to correctly stage score more than one-third of a group of moral thought statements while using a published self training rater guide. Further, a secondary analysis of the data (Napier, 1977) suggested that the reason why these elementary school teachers could not correctly stage score was that they were influenced by the content (choice and concepts used) of moral thought statements. However, there were some design problems with this prior experiment. First, the measure used to determine stage scoring ability had a low estimate of reliability. Second, the subjects used may have been unlike other elementary school teachers and, especially, unlike secondary school teachers. Third, the one group pretest-posttest experimental design did not allow for complete control over extraneous variables.

Two experiments were conducted to replicate the original study and examine the question of whether teachers can correctly stage score moral thought statements while using a rater guide. The second purpose of the two experiments was to replicate the secondary investigation and study the question of whether the content of moral thought statements influenced teachers as they stage scored.

Stage Scoring Moral Thought Statements

The process of stage scoring moral thought required a rater to ignore the content of moral thought statements as such and stage score on the basis of the structure of moral thought. The content of moral thought statements represented the choice made to do or not to do a moral act and the concepts used to justify the moral choice made. The structure of moral thought represented the way different concepts were used in justifying a moral decision. The different groups of similar

concepts used in moral thought statements were originally termed Aspects by Kohlberg, and the original stage scoring system was called Aspect scoring (Kohlberg, 1976).

For example, both a stage 1 response and a stage 5 response to a moral dilemma might contain the same choice to do a moral act as well as the same concepts of punishment, negative reactions, and condemnation (termed the Aspects of "Sanctions and Motives" by Kohlberg and the Aspects of "Motives for Engaging in Moral Action" by Rest). Although the content of the moral responses would be the same, the way the concepts were used to justify the choice made would be different. For a stage 1 response, the concepts of punishment, negative reactions, and condemnation would not be differentiated from the act itself (i.e., the act is always labeled as punishable, receiving negative reactions, and condemned by an external agent). For a stage 5 response, the same concepts would be differentiated to an internal judgment of the self in respect to contractual arrangements with the social group. It is this "qualitative" difference in the use of concepts which distinguished a stage classification of a moral thought statement as being either stage 1 or 5.

Originally, Kohlberg (1958) developed two rater manuals to aid in stage scoring moral thought statements. Kohlberg termed one a sentence coding guide and the other a global coding guide. The method of stage scoring used with the sentence guide required a rater to examine each moral thought content unit within the responses to all the dilemmas in a Moral Judgment Interview. These isolated moral thought content units varied in number, length, and content. The method of stage scoring used with the global guide required a rater to examine the entire response to each of the moral dilemmas used in a Moral Judgment Interview. This larger dilemma unit varied in length from a few sentences to several paragraphs, and like the moral thought content units, the larger dilemma units varied in content. Later, Porter and Taylor (1972) published a new global coding guide. This newer guide was a simplified version of Kohlberg's original global guide, and the first published rater guide

available to aid in stage scoring moral thought statements.

Besides the use of a rater guide, some form of training was usually required to stage score properly. Exact information on the nature of the training sequence used with the two original rater guides could not be found. Inferences made from the present training sequence offered by the Harvard Center for Moral Education indicated that the pattern of training most likely included: a) background information on the theory of and program for moral development; b) explanation of the rater guide and proper stage scoring; and c) practice with the rater guide. The Porter and Taylor rater guide did not have a specific training sequence. The Porter and Taylor guide was designed to be self training because the guide included background information on Kohlberg's theory of moral development as well as an explanation of the process of stage scoring.

One problem with these procedures for stage scoring was that a rater might not be able to stage score moral thought statements which were too far above the rater's own stage of moral thought development (Rest, Kohlberg, & Turiel, 1969; Rest, 1973). A rater who was validly stage scoring on the basis of structure would, however, correctly stage score any content within a comprehended stage. A rater who was validly stage scoring on the basis of structure might by chance incorrectly stage score moral thought statements within a stage which the rater comprehended, or might by chance correctly stage score moral thought statements within a stage the rater did not comprehend. However, a rater who was invalidly stage scoring on the basis of content would correctly and incorrectly stage score different contents within different stages in a non-random fashion because certain contents of moral thought statements would meet the rater's preconceived notion of appropriate choice and concepts for a particular stage no matter whether the rater comprehended a given stage or not.

Research Questions

The original study (Napier, 1976) examined only the Porter and Taylor self training rater guide in trying to determine whether teachers can stage score moral thought statements when using a rater guide. The present

study examined all three rater guides associated with the Aspect scoring system in studying the question of whether teachers can stage score moral thought statements. Experiment 1 compared the two original rater guides and Experiment 2 compared the combination of the two original guides with the Porter and Taylor rater guide while examining the question of teacher ability to stage score moral thought.

The question of whether the content of moral thought statements influenced teachers as they stage scored was examined in both experiments. In both experiments for each rater guide treatment used, the results of the test used were examined to determine whether teachers were incorrectly stage scoring each content within each stage randomly.

Dependent Measure

The instrument used to measure stage scoring ability, termed Moral Knowledge Test, has been described in detail in the previous study (Napier, 1976). The important characteristic of the instrument was that there were four different contents which crossed all six stages measured on the test (24 items). Content 1 consisted of the choice "Do" and the Aspects of "Orientation to Intentions and Consequences"; Content 2 consisted of the choice "Don't" and the Aspects of "Orientation to Intentions and Consequences"; Content 3 consisted of the choice "Do" and the Aspects of "Motives for Engaging in Moral Action"; and Content 4 consisted of the choice "Don't" and the Aspects of "Motives for Engaging in Moral Action." In this study a reliability coefficient of .81 (n = 72) was obtained on the Moral Knowledge Test using a Cronback alpha projected to a standard test (100 items).

Experiment 1

Procedures

Thirty-two preservice social studies teachers enrolled in a required five hour undergraduate curriculum course fall quarter were randomly assigned to two treatment groups. One treatment group was given the original sentence rater guide, and the other group was given the original global rater guide. Prior to distribution of the rater guide treatments, the

preservice teachers were given background information on Kohlberg's theory of and program for moral development. The preservice teachers were first given experiences in answering probe questions to four moral dilemmas. The purpose of this first exercise was to give the preservice teachers concrete experiences in what constitutes a moral dilemma. Next, the preservice teachers were given two readings on moral development and moral education (Kohlberg, 1971, 1975) and led in discussions of the theory and education programs for five class hours.

After the preservice teachers were given the rater guides, they were told to read the instruction on the use of the guide and to stage score examples of moral thought statements given as a homework assignment. They were told to do this homework without conferring with other classmates. Then the two groups met separately to discuss the homework assignment. During this class session the preservice teachers were led through the rater guide for their group with an emphasis on how to stage score validly on the basis of structure. At the end of this last training session, the preservice teachers were given the Moral Knowledge Test and instructed to classify the twenty-four statements using their respective rater guides as homework. Again they were told to complete the assignment independently. Two days later at the next class session the preservice teachers returned the Moral Knowledge Test. At this class session, the preservice teachers were interviewed to determine whether they had used the rater guides given. It was judged that the preservice teachers had tried to use the rater guides as they attempted to stage score the moral thought statements on the Moral Knowledge Test.

Results

The mean correct scores for each treatment group on each content within each stage, on each stage, on each content, and overall are presented in Table 1. The overall means for the two groups were almost the same. The analysis of variance test for repeated measures conducted on the Moral Knowledge Test correct scores (Table 2) confirmed this observation. There was no significant difference between the two original rater guide treatments

Table 1

Mean Scores on Moral Knowledge Test for Sentence Rater Guide and Global Rater Guide Treatments

Stage	Content 1	Content 2	Content 3	Content 4	Total
		Sentence Rater Guide Treatment			
1	0.063	0.375	0.563	0.938	1.939
2	0.563	0.125	0.625	0.438	1.751
3	0.875	0.375	0.688	0.688	2.626
4	0.500	0.688	0.688	0.688	2.564
5	0.500	0.750	0.125	0.188	1.563
6	0.750	0.625	0.250	0.250	1.875
Total	3.251	2.938	2.939	3.190	12.318
		Global Rater Guide Treatment			
1	0.188	0.688	0.750	0.938	2.564
2	0.438	0.188	0.500	0.563	1.689
3	1.000	0.375	0.875	0.500	2.750
4	0.500	0.875	0.563	0.375	2.313
5	0.625	0.500	0.313	0.250	1.688
6	0.438	0.563	0.063	0.188	1.252
Total	3.189	3.189	3.064	2.814	12.256

Note: Content 1 refers to the choice "Do" and Aspects of "Orientation to Intentions and Consequences"; Content 2 refers to the choice "Don't" and Aspects of "Orientation to Intentions and Consequences"; Content 3 refers to the choice "Do" and Aspects of "Motives for Engaging in Moral Action"; Content 4 refers to the choice "Don't" and Aspects of "Motives for Engaging in Moral Action".

Table 2

Analysis of Variance Test for Sentence Rater Guide and Global Rater Guide Treatments

Source	df	MS	F
Between Subjects	31		
Treatment	1	0.001	0.002
Subj w. groups	30	0.528	
Within Subjects	736		
Stage	5	1.799	8.695*
Treatment x Stage	5	0.351	1.698
Stage x subj w. groups	150	0.207	
Content	3	0.057	0.306
Treatment x Content	3	0.099	0.530
Content x subj w. groups	90	0.186	
Stage x Content	15	2.076	11.332*
Treatment x Stage x Content	15	0.232	1.266
Stage x Content x subj w. groups	450	0.183	
Total	767	0.250	

*Significant at $p < .05$

on overall correct stage scoring. Both rater guides only helped the group of preservice teachers correctly stage score 51% of the moral thought statements on the Moral Knowledge Test.

The mean scores for contents within each stage for both treatment groups (Table 1) indicated that the preservice teachers were not stage scoring each content within each stage equally. This observation was confirmed by the analysis of variance test (Table 2). The significant interaction between the factors of Stage and Content meant that the preservice teachers were correctly stage scoring the different contents within each stage non-randomly. The insignificant finding for the interaction between Treatment, Stage, and Content meant the different groups were correctly stage scoring the different contents within each stage in a similar fashion. This finding was interpreted to mean that the preservice teachers were being influenced by content while using both rater guides. The reason the preservice teachers did not correctly stage score more than 51% of the moral statements on the Moral Knowledge Test was because they were invalidly stage scoring on the basis of content.

Experiment 2

Procedures

Forty preservice social studies teachers enrolled in a required five hour undergraduate curriculum course winter quarter were randomly assigned to two treatment groups. One treatment group was given both the original sentence rater guide and global rater guide developed by Kohlberg, and the other group was given the Porter and Taylor global rater guide. The combined original rater guide represented a guide which permitted a rater to stage score once using the global rater guide and then cross validate using the sentence rater guide.

Prior to distribution of the two rater guides, the preservice teachers were given the same background information used in Experiment 1. It should be noted that although the Porter and Taylor rater guide was self training, the self training sequence was not used in this study. Instead the preservice teachers assigned to the Porter and Taylor rater guide received

the same training as the preservice teachers who used the combined original rater guide.

After the background information was given, these preservice teachers followed the same training with their respective rater guides and took the Moral Knowledge Test used in Experiment 1. Like the first experiment, informal interviews with the preservice teachers indicated that they had tried to use the rater guides given.

Results

The mean correct scores for each treatment group on each content within each stage, on each stage, on each content, and overall are presented in Table 3. The overall means for the two groups were again almost the same. The analysis of variance test for repeated measures conducted on the Moral Knowledge Test correct scores for Experiment 2 (Table 4) confirmed this observation. There was no significant difference between the two rater guide treatments on overall correct stage scoring. Both rater guides only helped the group of preservice teachers correctly stage score about 44% of the moral thought statements on the Moral Knowledge Test. This was a lower percentage than the finding of the first experiment and may have resulted from differences in the preservice teachers in the two experiments. Nevertheless, the rater guide treatments used did not help these preservice teachers correctly stage score most of the moral thought statements given.

The mean scores for contents within each stage for both treatment groups (Table 3) again indicated that the preservice teachers were not stage scoring each content within each stage equally. This observation was also confirmed by the analysis of variance test (Table 4). As found in the first experiment, the interaction between Stage and Content was significant. Unlike the first experiment, the interaction between Treatment, Stage and Content was also significant. This latter finding indicated that the preservice teachers were correctly stage scoring the different contents at each stage in a dissimilar fashion depending on the rater guide used. Nevertheless, the results of the analysis of variance test was again interpreted to mean that the preservice teachers were

Table 3

Mean Scores on Moral Knowledge Test for Original Rater Guide and Porter and Taylor Rater Guide Treatments

Stage	Content 1	Content 2	Content 3	Content 4	Total
	Original Rater Guide Treatment				
1	0.350	0.850	0.450	0.900	2.550
2	0.600	0.050	0.650	0.450	1.750
3	0.800	0.350	0.550	0.300	2.000
4	0.350	0.550	0.500	0.100	1.500
5	0.650	0.850	0.050	0.150	1.700
6	0.650	0.300	0.100	0.050	1.100
Total	3.400	2.950	2.300	1.950	10.600
	Porter and Taylor Rater Guide Treatment				
1	0.050	0.650	0.550	0.900	2.150
2	0.350	0.400	0.650	0.600	2.000
3	0.750	0.200	0.700	0.500	2.150
4	0.150	0.650	0.150	0.400	1.350
5	0.500	0.250	0.300	0.150	1.200
6	0.600	0.400	0.300	0.250	1.550
Total	2.400	2.550	2.650	2.800	10.400

Note: Content 1 refers to the choice "Do" and Aspects of "Orientation to Intentions and Consequences"; Content 2 refers to the choice "Don't" and Aspects of "Orientation to Intentions and Consequences"; Content 3 refers to the choice "Do" and Aspects of "Motives for Engaging in Moral Action"; Content 4 refers to the choice "Don't" and Aspects of "Motives for Engaging in Moral Action".

Table 4

Analysis of Variance Test for Original Rater Guide and Porter and Taylor Rater Guide Treatments

Source	df	MS	F
Between Subjects	39		
Treatment	1	0.017	0.061
Subj w. groups	38	0.217	
Within Subjects	920		
Stage	5	1.715	7.396*
Treatment x Stage	5	0.357	1.538
Stage x subj w. groups	190	0.232	
Content	3	0.392	2.424
Treatment x Content	3	1.108	6.861*
Content x subj w. groups	114	0.162	
Stage x Content	15	2.343	12.577*
Treatment x Stage x Content	15	0.482	2.585*
Stage x Content x subj w. groups	570	0.186	
Total	959	0.246	

*Significant at $p < .05$

being influenced by content while using both rater guides. So, the reason the preservice teacher did not correctly stage score more than 44% of the moral thought statements was because they were invalidly stage scoring on the basis of content.

Conclusion

The original study found no relationship between age and teaching experience with stage scoring ability. Therefore, the conclusions of this present study based on preservice social studies teachers should generalize to experienced social studies teachers. The findings of the two experiments support the conclusions made in the original study that teachers cannot adequately stage score moral thought statements with the aid of a training sequence and a rater guide. Furthermore, the findings support the conclusions that the reason teachers cannot adequately stage score is because they invalidly stage score on the basis of the content of moral thought.

The training sequence used in the two experiments was similar to that presently used at the Harvard Center for Moral Education. Kohlberg and his associates did note that their training sessions would not make perfect stage scorers, but the training sequence should make valid stage scorers. Perhaps teachers need a different training sequence before they can validly stage score.

The rater guides used in the two experiments were associated with the Aspect scoring system. The findings of this and the previous studies empirically confirmed the intuitive judgment of Kohlberg and his associates that the Aspect scoring system and related rater guides were susceptible to stage scoring invalidly on the basis of content (Kohlberg, 1973; 1976). Kohlberg and his associates are developing a new scoring system (Issue scoring) and an accompanying rater guide (Kohlberg, Colby, Gibbs, Speicher-Dubin, & Power, 1976). Perhaps the newer scoring system and related guide would help teachers validly stage score.

Until research is done on different training sequences as well as the newer scoring system and rater guide, social studies teacher educators

should continue to stress the findings of this and the previous studies. At present, teachers should not try to stage score moral thought statements because they most likely stage score invalidly on the basis of content.

Footnotes

[1]I have assumed most social studies educators are familiar with Kohlberg's theory of moral development which he has described as having three major levels with two stages within each major level. The Preconventional level contained the Punishment and obedience orientation stage (1) and the Instrument relativist orientation stage (2). The Conventional level contained the "Good boy-nice girl" orientation stage (3) and the "Law and order" orientation stage (4). The final level was termed Post-conventional and consisted of the Social-contract legalistic orientation stage (5) and the Universal ethical principle orientation stage (6).

[2]Fenton and Kohlberg seem to contradict themselves. First they claim that Kohlberg's measurement system is too complicated for teachers to use for the purpose of evaluating changes in students' moral stages. Then Fenton and Kohlberg claim that teachers can stage score for the purpose of supplying +1 models. Intuitive stage scoring to supply +1 models is more difficult than using a rater guide to evaluate student changes. Fenton and Kohlberg need to resolve this contradiction.

References

Blatt, M. & Kohlberg, L. The effects of classroom moral discussion upon children's level of moral judgment. Journal of Moral Education. 1975, 4, 129-161.

Fenton, E. & Kohlberg, L. (Eds.) Teacher training in values education: A workshop. Pleasantville, N. J.: Guidance Associates Filmstrips, 1976 (a).

Fenton E. & Kohlberg, L. (Eds.) The values in a democracy series. Pleasantville, N. J.: Guidance Associated Filmstrips, 1976 (b).

Galbraith, R. & Jones, T. Moral reasoning: A teaching handbook for adapting Kohlberg to the classroom. Anoka, Minn.: Greenhaven Press, 1976.

Kohlberg, L. The development of modes of moral thinking and choice in the years ten to sixteen. Unpublished Ph.D. dissertation, University of Chicago, 1958.

Kohlberg, L. The concepts of developmental psychology as the central guide to education: Examples from cognitive, moral, and psychological education. In M. Reynolds (Ed.), Proceedings of the conference on psychology and the process of schooling in the next decade: Alternative conceptions. Minneapolis: University of Minnesota Department of Audio-Visual Extention, 1971.

Kohlberg, L. Continuities in childhood and adult moral development revisited. In P. Baltes & K. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic Press, 1973.

Kohlberg, L. The cognitive-developmental approach to moral education. Phi Delta Kappan. 1975, 41, 670-677.

Kohlberg, L. Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), Moral development and behavior: Theory, research, and social issues. New York: Holt, Rinehart & Winston, 1976.

Kohlberg, L., Colby, A., Gibbs, J., Speicher-Dubin, B. & Power, C. Assessing moral stages: A manual. Cambridge: Harvard Center for Moral Education, 1976.

Napier, J. The ability of elementary school teachers to stage score moral thought statements. Theory and Research in Social Education. 1976, 4(2), 39-56.

Napier, J. Content influence while stage scoring moral thought statements. Manuscript submitted for publication, 1977.

Porter, N. & Taylor, N. How to assess the moral reasoning of students. Toronto: Ontario Institute for Studies in Education, 1972.

Rest, J. The hierarchial nature of moral judgment: A study of patterns of comprehension and preference of moral stages. Journal of Personality. 1973, 41, 86-109.

Rest, J. Developmental psychology as a guide to value education: A review of "Kohlbergian" programs. Review of Educational Research. 1974, 44, 241-259.

Rest, J., Kohlberg, L. & Turiel, E. Level of moral development as a determinant of preference and comprehension of moral judgments made by others. Journal of Personality. 1969, 37, 255-252.

Turiel, E. Developmental processes in the child's moral thinking. In P. Mussen (Ed.), Trends and issues in developmental psychology. New York: Holt, Rinehart & Winston, 1969.